

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
8 February 2001 (08.02.2001)

PCT

(10) International Publication Number  
**WO 01/09785 A1**

(51) International Patent Classification<sup>7</sup>: **G06F 17/60**

(21) International Application Number: **PCT/KR00/00836**

(22) International Filing Date: **31 July 2000 (31.07.2000)**

(25) Filing Language: **Korean**

(26) Publication Language: **English**

(30) Priority Data:  
1999/31624 31 July 1999 (31.07.1999) KR  
2000/18398 8 April 2000 (08.04.2000) KR

(71) Applicant and

(72) Inventor: **PARK, Kyu, Jin** [KR/KR]; 103-1206  
Dukpoong Hyundai Apts., 524, Dukpoong-dong,  
Hanam-si, Kyungki-do 465-010 (KR).

(74) Agent: **LEE, Wan, Hwee**; #407 Sungjee Heights 3 Building,  
642-6, Yeoksam-dong, Kangnam-ku, Seoul 135-080 (KR).

(81) Designated States (*national*): AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

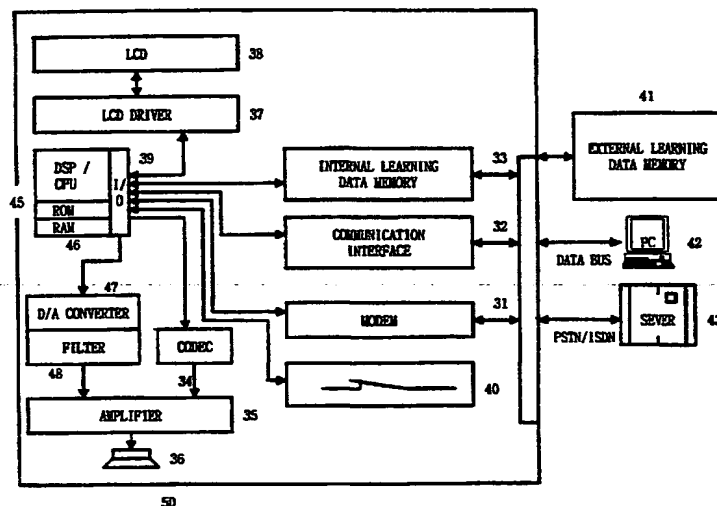
(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**

— *With international search report.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: **STUDY METHOD AND APPARATUS USING DIGITAL AUDIO AND CAPTION DATA**



(57) Abstract: A method and an apparatus for learning by using a digital audio and its synchronized caption data is disclosed. When learning a language or a song, the voice, the melody accompaniment and the text can be simultaneously or selectively outputted. This can be realized in a reproduction apparatus which is capable of storing the digital audio files and the caption data. The outputting of the voice, the melody accompaniment and the caption data can be adjusted in accordance with the progress of the learning. The reproduction apparatus should have two or more channels, and the channels can store different contents. The different channels can be arbitrarily selected by the user.

**THIS PAGE BLANK (USPTO)**

## STUDY METHOD AND APPARATUS USING DIGITAL AUDIO AND CAPTION DATA

### FIELD OF THE INVENTION

5       The present invention relates to a method and an apparatus  
for learning by using a digital audio and its synchronized caption  
data. More specifically, the present invention relates to a method  
and an apparatus for learning by using a digital audio and the  
selection of the output channel for its synchronized caption data, in  
10       which in the case where particular subjects such as foreign  
language, song words, melodies and the like are needed be learned  
repeatedly, the learning is carried out by adjusting the difficulty  
levels in accordance with the learner's progress of the learning, so  
that a self-learning may be possible.

### BACKGROUND OF THE INVENTION

15       In accordance with the progress in the digital signal  
processing technology, various products which utilizes digital audio  
signals are developed and sold. The examples are the MP3 player,  
the language learning apparatus using a digital audio file, and the  
karaoke for outputting the melody accompaniment by utilizing a  
20       digital audio file. Such apparatus outputs not only the song words  
and melody accompaniment but also caption data in letters. That  
is, together with the outputting of the audio signals, letters are  
displayed, and thus, they are helpful in the language learning and  
in the song learning.

25       Essentially, the digital audio data includes only vocal  
information. However, this digital audio data can store a caption

information. When playing the digital data, the output can be obtained through a voice outputting device such as earphone and through a display device such as LCD.

5 The bit arrangement of the digital audio data consists of frames or AAU (audio access unit). These frame units cover the MP3 apparatus and the audio parts of all the DVD (digital versatile disk) standard and the MPEG standard.

10 The software which is capable of inserting the caption data into the digital audio data can express the caption display position by the frame numbers, and therefore, it can be applied to all the digital audio data in which the bit stream is arranged in the form of frame units.

15 However, in the conventional language learning apparatuses for reproducing the digital audio data, the user can only unilaterally listen and watch the outputted digital audio data and the caption data, the former being outputted through a speaker or an earphone. Therefore, the learner cannot set diversified situations, and therefore, the learning of a language cannot be effective.

20 Further, in the case of karaoke, the song words and the melody accompaniment are simultaneously outputted, and therefore, a person who does not know can sing the song by watching the displayed letters. However, in this case also, the person should roughly know the song words beforehand. That is, if the person  
25 is to sing the song well, then the person has to have been familiar to the song of the original singer.

30 Accordingly, there has come a demand for an apparatus in which the conventional audio apparatus and the karaoke are combined together in such a manner that the voice of the original singer can be selectively outputted.

## SUMMARY OF THE INVENTION

The present invention is intended to facilitate the learning of languages and songs.

Therefore it is an object of the present invention to provide  
5 a learning method utilizing a digital audio and its caption data, in which the difficulty level is adjusted in accordance with the progress of the learning, so that the learning would be facilitated, and that the learning can be carried out for oneself.

It is another object of the present invention to provide a  
10 learning method and a learning apparatus utilizing a digital audio and the output channel selection for its caption data, in which in the case of a language learning, the user can selectively set the audio outputting situation, so that the user can perform the desired role.

15 In achieving the above objects, the method for learning by using a digital audio and its caption data according to the present invention includes the steps of: forming a first learning pattern storing mode for storing a song caption, the voice of an original singer, and a melody accompaniment by converting their signals  
20 into a digital file; and forming a second learning pattern storing mode for storing a song caption and a melody accompaniment by converting their signals into a digital file, whereby a digital file is formed for an arbitrary song, and the digital file is reproduced based on the first or second learning pattern storing mode so as to  
25 facilitate learning an arbitrary song.

In another aspect of the present invention, the method for learning by using a digital audio and its caption data according to the present invention includes the steps of: forming a first learning pattern storing mode for storing a foreign language speech or a

news by distinguishing the voice of a speaker and the caption of speech details in letters or news details in letters, and by converting signals of the audio and caption to a digital file; and forming a second learning pattern storing mode for storing a  
5 foreign language speech or a news by distinguishing the voice of a speaker and the caption of speech details in letters or news details in letters, and by converting the signals of only the voice of the speaker to a digital file, whereby a digital file is formed for an arbitrary speech or news, and the digital file is reproduced in  
10 accordance with a selection of a reproduction by the user so as to make it possible to learn an arbitrary speech or news.

In still another aspect of the present invention, the method for learning by using a digital audio and its caption data according to the present invention includes the steps of: forming a first  
15 learning pattern storing mode for recording a full sound - full caption by preparing a digital data file of all the voices and all the talk captions of all talkers of a foreign movie; and forming a second learning pattern storing mode for storing a data file by recording a scenario of the movie after deleting the voices of  
20 certain talkers so as to make a user talk in place of the deleted voices, whereby a digital data is formed, and if the user selects a learning reproduction mode and selects the talkers, the digital data file is selectively reproduced so as to make the user talk in place of particular talkers.

25 In still another aspect of the present invention, the method for learning by using an output channel selection for a caption data according to the present invention includes the steps of: checking the operation mode of a current reproduction operation upon inputting an operation-on signal by the user for reproducing  
30 audio signals (first step); outputting audio signals which have been

set to respective channels (R and L) if the operation mode is found to be a normal channel outputting (second step); reproducing and outputting the audio signals to the right channel if the operation mode is set to the right channel (R) (third step); and reproducing and outputting the audio signals to the left channel (L) if the operation mode is set to the left channel (fourth step).

In still another aspect of the present invention, the learning apparatus for learning by using an output channel selection for a caption data according to the present invention is characterized in that: if an operation-on signal for reproducing audio signals from a keypad is an input, the operation mode during a reproduction which is currently set by a control section is checked; if the operation mode is normal, the control section controls a decoder to output the audio signals which have been set to respective channels (R and L); if the operation mode is set to the right channel (R), the control section controls the decoder to reproduce and output the audio signals to the right channel; and if the operation mode is set to the left channel (L), the control section controls the decoder to reproduce and output the audio signals to the left channel.

## BRIEF DESCRIPTION OF THE DRAWINGS

The above objects and other advantages of the present invention will become more apparent by describing in detail the preferred embodiment of the present invention with reference to the attached drawings in which:

FIG. 1 is a block diagram showing the constitution of the digital audio player as an example of the hardware which is applied to the learning method according to the present invention;

FIG. 2 is a flow chart showing the input/output procedure for the digital audio and its caption data in the learning method for learning the songs according to the present invention;

5 FIG. 3 is a flow chart showing the input/output procedure for the digital audio and its caption data in the learning method for learning the foreign language speeches according to the present invention;

10 FIG. 4 is a flow chart showing the input/output procedure for the digital audio and its caption data in the learning method for learning the foreign languages through foreign movie scenarios and their sound tracks according to the present invention;

FIGs. 5a to 5c illustrate the output status of the caption picture for the respective learning foreign movies;

15 FIG. 6 is a partial block diagram showing a conventional stereo reproducing apparatus which is an example of the hardware used in the present invention;

FIG. 7 is a partial block diagram showing a conventional multi-channel reproducing apparatus which is an example of the hardware used in the present invention;

20 FIG. 8 is a flow chart showing the learning method utilizing the output channel selection for the caption data with the stereo channel adopted according to the present invention;

25 FIG. 9 is a flow chart showing the learning method utilizing the output channel selection for the caption data with the multiple channels adopted according to the present invention; and

FIG. 10 illustrates the constitution of a personal computer in which the learning method using the output channel selection is adopted for the caption data according to the present invention.

#### DETAILED DESCRIPTION OF THE PREFERRED



## EMBODIMENTS

The preferred embodiments of the present invention will be described referring to the attached drawings.

### Example 1

5       The learning method utilizing a digital audio and its caption data according to the present invention includes: (1) a method of selectively selecting the output status by making the digital audio storing mode and the caption data storing mode different from each other; and (2) a method of selectively setting the output  
10       status after storing the digital audio and the caption data in different channels (more than stereo channels).

      In the present invention, the former and latter are distinguished into: a method in which a digital audio and caption data are utilized, and a method in which the digital audio and an  
15       output channel selection for the caption data are utilized. In principle, it is apparent that there is a similarity between the two methods of the present invention, in that the digital audio and the caption data are utilized in the learning. First, referring to FIGs. 1 to 5, that is, in Examples 1 to 3, the method of utilizing the  
20       digital audio and the caption data will be described, and then, referring to FIGs. 6 to 10, that is, in Example 4, the method of utilizing the digital audio and the output channel selection for the caption data will be described.

      FIG. 1 is a block diagram showing the constitution of the digital audio player as an example of the hardware which is  
25       applied to the learning method according to the present invention.

      As shown in this drawing, the digital audio player 50

includes: a modem 31 for receiving a caption digital data from a caption learning network server 43 of a wired switching station through a PSTN/ISDN network; a communication interface 32 for receiving a readable data by an internal device through a data bus from a PC 42 based on the transmission data; and an internal on-screen letter language learning data memory 33 for storing the language learning voices and caption data, the memory 33 being connected through a connector 44 to an external learning data memory 41.

The modem 31, the communication interface 32 and the internal learning data memory 33 are connected to a DSP/CPU 39 which has an I/O port, a ROM 45 and a RAM 46.

The DSP/CPU 39 is connected to a switch having PLAY, REW, FF and STOP keys, and is also connected to an LCD 38 which displays the caption data after converting it into letters. The digital audio signals which has been processed by the DSP/CPU 39 are transferred through a CODEC 34, a converter 47 and a filter 48 to be finally outputted through a voice output device 36.

When the digital audio player receives the caption language learning data from an external device, the data source is a data base server 43 of a wire switching station to form a modem communication mode, and the CPU is connected to a server of the wire switching station, and that is, the modem 31 is driven to carry out a DTMF dialing.

Further, in the digital audio player, the required digital data can be received from the PC 42 through the interface device 32.

The DSP/CPU 39 processes the digital audio and caption data after receiving them from the modem 31 or from the

communication interface 32 to store them into the internal learning data memory 33.

5 The communication interface 32 is connected through a wire device such as a computer (parallel) printer port, a serial port, a USB, or a firewire (IEEE 1394), or through a wireless form such as an infrared ray data or a blue tooth, so that the data can be stored into the storing means of the reproduction apparatus, i. e., to the learning data memory 33. The storing means may be a non-volatile memory such as a flash memory, or a read/write  
10 storing means such as a DVD (digital versatile disk).

The switching section 40 which is connected to the DSP/CPU 39 selects various functions of the digital audio player. For example, if the PLAY switch of the switching section 40 is turned on, the CPU 39 puts the player to a learning reproduction  
15 mode, and brings the selected digital file from the internal learning data memory 33 to process it.

The digital audio data which has been processed by the DSP/CPU 39 is outputted in an analogue voice after transferring the signals through the CODEC 34, the converter 47 and the filter  
20 48. Meanwhile, the caption data which has been processed by the DSP/CPU 39 is displayed on an LCD 38 after passing an LCD driver 37.

In this manner, through the simultaneous outputting of the voice and letters, there can be improved the language learning  
25 effect.

FIG. 2 is a flow chart showing the input/output procedure for the digital audio and its caption data in the learning method for learning the songs according to the present invention.

As shown in this drawing, if a song is selected, then there  
30 is checked as to whether it is a song digital data file preparation

mode. If it is the relevant mode, then at a first learning pattern storing mode, a distinguishing is carried out into the voice of the original singer, the melody accompaniment and the word caption. In this manner, a digital data file is formed and recorded like the  
5 karaoke.

In the above, the subject was popular songs, but as long as there are present the song words, the voices of the original singers and the melody accompaniments, or as long as there are present the song words and the voices of the original singer, then  
10 any kinds of songs such as classics, semi-classics, children's songs and the like can be adopted. In this context, the songs which are mentioned below should be understood to be all kinds of songs.

Then at a second learning pattern storing mode, a digital  
15 data is prepared by employing only the melody accompaniment and the caption data. Under this condition, a judgment is made as to whether the song consists of voices of duet singers. If not, then at a third learning pattern storing mode, a digital data file is prepared by employing only the melody accompaniment. Under  
20 this circumstance, the third learning pattern storing mode can be skipped.

However, after carrying out the second learning pattern storing mode, if the song is found to be the voices of duet singers, then at a fourth learning pattern storing mode, a digital file is  
25 prepared by adopting only the voice and caption data of a singer a. Then at a fifth learning pattern storing mode, a digital file is prepared by adopting only the voice and caption data of a singer b.

Such separate storing of the songs can be done for as  
30 many songs as desired.

Thereafter, if the user executes the desired songs through a selection of the reproduction mode, then play of the relevant songs can be carried out, with the result that the language learning becomes interesting.

5           For example, if one is familiar with the song words, then only the melody accompaniment can be outputted. Or one can selects the second learning pattern storing mode, and can exercise the song without watching the song words. If one is not good with both the song words and the melody, then both the song  
10 words and the melody accompaniment can be simultaneously outputted. In this manner, the option of the user is arbitrary.

#### Example 2

FIG. 3 is a flow chart showing the input/output procedure for the digital audio and its caption data in the learning method  
15 for learning the foreign language speeches or news according to the present invention.

Here, if the user selects a speech or a news for learning a language, then the system judges as to whether it is a language learning digital data file preparing mode using a speech or a news.  
20 If it is the digital preparing mode, that is, if it is the learning data inputting mode, then at a first learning pattern storing mode, a digital data file is formed by loading the caption data such as the speech or news together with the audio data of the speaker.

During a judgment as to whether a translation is required  
25 or not, if it is not required, then at a second learning pattern storing mode, only the voice of the speaker is loaded in the digital data file to record it.

If it is found that a double caption mode is present with a

simultaneous translation accompanied, then at a third learning pattern storing mode, the single LCD screen is divided into two areas when the voice of the speaker is outputted, so that one area of the LCD screen can display the original caption letters, and that  
5 the other area of the LCD screen can display the translated version of the original language. Then the prepared digital data file is recorded.

In this manner, the user can make a selection of speeches and news in accordance with the taste and the understanding  
10 ability. Therefore, a foreign language can be efficiently learned by adopting the speeches or news.

### Example 3

FIG. 4 is a flow chart showing the input/output procedure for the digital audio and its caption data in the learning method  
15 for learning the foreign languages through foreign movie scenarios and their sound tracks according to the present invention.

Here, a language learning pattern using a movie scenario and a real time sound track is selected. Then the system judges as to whether it is a digital data file preparing mode using a  
20 scenario and its sound track. If it is the digital file preparing mode, i.e., a learning data inputting mode, then at a first learning pattern storing mode, all the voices of the talkers of the movie, the names of talkers and the caption letters of the talkers are entered into a digital data file. Thus a full sound - full caption is  
25 recorded for the movie. In this case, the caption data is displayed on the LCD as shown in FIG. 5a.

Then at a second learning pattern storing mode, a full sound condition is carried out. That is, the caption data is

outputted, while the real time voice output is muted in storing the file. At this second learning pattern storing mode, the recording is carried out for each of the talkers separately.

Under this condition, the caption of the talkers can be displayed in a blinking form in a predetermined sequence.

At this second learning pattern storing mode, the user can speak in place of a certain talker by carrying out a dubbing mode. In this manner, the user can confirm the correctness of his or her own pronunciation, and if the pronunciation is insufficient or incorrect, then the user can correct his or her pronunciation.

For this purpose, the voices of the user can be fed back behind the voices of the original talkers through a mike of the digital audio signal processing apparatus. In this way, the user can listen to his or her own pronunciation.

At a third learning pattern storing mode, a digital data file is prepared as follows. That is, the name of a relevant talker is outputted, while the sound track audio and the caption data are muted and turned to a blank interval respectively. This requires a high memorizing ability, and therefore, its actual utility is very low. Therefore it may be deleted.

In this manner, the user participates in the foreign movie by taking the place of a talker of the movie, and therefore, the language learning efficiency can be improved.

Instead of the names of the talkers, there are assigned serial codes to the respective talkers, and each of the serial codes is matched to each of the relevant talkers. In this manner, each of the caption data for each of the talkers can be separately stored.

Thus a learning data base is constructed by using the scenario and the audio of a foreign movie. In this state, the user selects the talker for whom the user wants to talk instead of the

original talker. Then the desired talker can be sorted out to be separately outputted. Or a relevant talker can be deleted.

5 As shown in FIG. 5b, if the user wants to take part in the movie by selecting a particular role, then his or her own voices are fed back into his or her own ears through the mike of the digital audio signal processing apparatus, because he or she has talked in place of the particular talker. The user can recognize any incorrectness of the pronunciation, so that the incorrect pronunciation can be corrected in learning the foreign language.

10 Further as shown in FIG. 5c, even when the particular talker deleting mode is executed, the name of the original talker can be unsuppressed, but can be made present, so that the user would feel as if the user were the real actor. Thus the sensation and feeling can be expressed in a natural manner, thereby making it possible to improve the efficiency of the learning the foreign language.

15 Through repetitions of this participated learning, the native pronunciation of the foreign language can be learned, thereby realizing a high learning efficiency. Further, depending on the selection by the user, the voices and the caption data of a particular talker can be deleted in the same manner, and therefore, the learning of the foreign language can be enhanced.

20 In the above, description were made that diversified learning data can be stored in the digital storing means of the digital audio player, and that the stored contents can be selectively read out to carry out the language learning by taking examples. However, it is also possible that various prepared data can be downloaded from a PC or from a data server to store them and to selectively read out them so as to carry out the learning of the foreign language.

25

30



#### Example 4

In this example, the selection of output channel for the caption data and the digital audio is adopted in learning a foreign language. Referring to FIGs. 6 and 7, first the conventional  
5 stereo-channel or multi-channel reproduction apparatus will be briefly described as to their operations. Then with reference to this, the present invention will be described.

The multi-channel reproduction apparatus (FIG. 7) which is related to the method of the present invention includes: an external  
10 data storing memory 110; an external interface 190 for transmitting and receiving the data to and from an external apparatus; a user input keypad 180; a control section 120 with a program installed therein for driving the overall system; a decoder 130 for converting digital audio signals; a DAC 140 for converting the  
15 converted analogue signals of the decoder 130 to output them through at least multiple channels to a speaker; and a screen driving device 160 for driving a picture display device 170, the picture display device 170 displaying the caption data.

The memory 110 is a storing means for storing the digital  
20 audio file data after receipt of it from an external source. The stored audio file can be reproduced by the control signals of the user. The audio file either has been stored during the manufacture of the product before the selling of the product, or can be stored after the manufacture by downloading the audio file from a PC or  
25 other external source through the external interface 190. The external interface 190 will be described later. The caption data is also stored in the memory 110, and the caption data is read out from the memory during the reproduction. For example, the memory 110 may be a non-volatile memory such as a flash

memory or an optical disk such as DVD, while other kinds of storing means may be usable. The memory 110 is detachably or fixedly installed within the reproduction apparatus.

5 The keypad 180 is for inputting commands for reproduction of the audio file, and includes a recording key, a reproduction key, a mode selection key and the like. That is, the keypad 180 includes functions keys such as a reproduction function key, a repeated reproduction key, a mode selection key (normal, and left and right channels). The control signals which are inputted by  
10 the user are inputted through the keypad 180 to the control section 120.

The control section 120 consists of a microcomputer, and is stored with a program for executing the reproduction and the caption display. Further, the control section 120 is connected to  
15 the interface 190, for receiving digital files from an external source. The control section 120 further stores a program for outputting the caption data to the picture display device in synchronization with the output of the audio signals.

The interface 190 can be variously constituted such that it  
20 can transmit the data through a wire such as a printer port (parallel port), a serial port, USB (universal serial bus), firewire (IEEE 1394) or the like, or through a wireless route such as blue tooth.

The control section 120 is connected to the decoder 130 for  
25 converting the digital audio signals. The decoder 130 converts the stored audio signals which have been recorded through the multiple channels. For example, the decoder 130 can be constituted by using the chips such as AAC, AC-3 or the like which can reproduce the various multi-channel digital audio  
30 signals.

The digital audio signals which have been converted by the decoder 130 are digital signals, and therefore, they are reconverted to analogue audio signals by the DAC 140. The outputted signals are outputted the speakers 150 and 152 for the channels

5 respectively, thereby realizing a sound mixing effect.

FIG. 6 shows two speakers, but their number can be increased or decreased depending on the number of the channels which are assigned to the decoder 130. FIG. 7 illustrates a plurality of speakers based on the multi-channel method. Further,  
10 the present invention can be applied to the case where a headphone or an earphone is used like in the conventional method. All these should come within the scope of the present invention.

Reference code 160 is an on-screen caption driving device which is operated by the control signals of the control section 120.  
15 Reference code 170 is a picture display device for displaying the caption data by being activated by the picture driving device 160. This picture display device may be an LCD or a CRT. If an audio is reproduced, the control section 120 outputs the caption data which is synchronized to the audio, the outputting being done  
20 through the picture display device 170. Thus the audio signals are outputted through the speaker, while the synchronized caption data is displayed on the picture display device 170. Thus the user can learn the language while watching the caption data and listening to the audio output.

25 The size of the caption data block is decided in view of the size of the picture display device 170, and the respective caption data blocks are synchronized with the audio outputs. That is, the audio signals have the information on the starting position of each of the caption data block.

30 The control section 120 outputs the caption data to the

picture display device 170 in synchronization with the audio signals by utilizing the above mentioned position information. That is, the control section 120 monitors the position information in the audio signals which are being reproduced. Then the control  
5 section 120 compares the position information of the audio signals with the position information of the caption data. Then at the instant a synchronization occurs, the caption data is displayed on the picture display device 170.

With the above described apparatus, the learning method  
10 according to the present invention includes the steps of: checking an operation mode of a currently set reproduction by a control section 120 upon inputting an operation signal for reproduction of audio signals through a keypad 180; controlling a decoder 130 by the control section 120 (if the operation mode is normal) to output  
15 the audio signals to respective right and left channels (R and L); reproducing and outputting the audio signals by the control section 120 to the right channel R by controlling the decoder 130 if the operation mode has been set to the right channel R; and reproducing and outputting the audio signals by the control section  
20 120 to the left channel L by controlling the decoder 130 if the operation mode has been set to the left channel L.

For the sake of describing convenience, the digital audio file is assumed to be a stereo file in which two channels are present as shown in FIG. 6. However, in the multi-channel  
25 recording method, a greater plurality of channels are provided, in such a manner that the channels can be controlled separately for each of them. In the stereo channels, there are provided the normal, left and right channels only, while in the multi-channel method, the number of the channels are increased.

30 The user selects one mode from among the normal mode,

the left channel outputting mode and the right channel outputting mode by pressing the relevant function key of the keypad 180. After selection of the function key, the user selects the language learning data or the karaoke song, so that the selected one would  
5 be inputted into the control section 120 by pressing the reproduction key of the keypad 180. When the reproduction is started, the caption data is displayed on the picture display device 170 in synchronization with the audio signals.

The control section 120 first checks the status of the  
10 setting of the operation mode before shifting to the reproduction mode. This is stored in the internal memories (RAM and ROM) of the control section 120, and when needed, it is brought out to be used. As a result of checking the operation mode, if it is found to be the normal mode, then the control section 120 outputs a control  
15 signal to the decoder 130 to reproduce the relevant audio file, so that the audio signals would be outputted through the left and right channels to the speakers 150 and 152. Thus the two speakers 150 and 152 outputs the audio signals simultaneously after receipt of them through the left and right channels. At the  
20 same time, the caption data which is synchronized with the audio signals is displayed to the picture display device 170, and therefore, the user can learn the language by listening to the audio output while watching the caption data.

Meanwhile, if the operation mode is found to be a left  
25 channel mode or a right channel mode, then the control section 120 outputs a control signal to the decoder 130 so that the signals of only the relevant channel would be outputted. The decoder 130 decodes only the signals of the relevant channel, and the outputted digital audio signals are converted to analogue signals by the DAC  
30 140 to be finally outputted through one of the speakers 150 and

152.

This outputting will be described based on an example.

In the case of the language learning, it is assumed that there are two talkers A and B, and that the talks between them  
5 are respectively stored into the left and right channels. If the user wants to learn the language by memorizing the talks of the talker A, and wants to talk with the talker B, then the channel in which the talks of the talker B is turned on, while the channel of the talker A is turned off. That is, the operation mode is set in this  
10 way.

After setting the operation mode in this way, if the user activates the reproduction apparatus, then the channel of the talker A is muted all the time. Thus the user can carry out the language learning after memorizing the letters or by watching the  
15 displayed caption data. The caption data has not been subjected to any selection mode, and therefore, the caption data is displayed in the normal manner. However, the caption data can also be subjected to a selection mode, to selectively output it.

Further, a selected channel and non-selected other channels  
20 can be activated simultaneously. The reason is as follows. If a single channel is turned on to output the signals only through the selected channel, then the rest of the channels are inactivated, and only one speaker is activated. That is, the audio signals are outputted through only the single speaker, and this may give the  
25 reality feeling. However, if the user listens through only one speaker or through only one earphone, then the hearing balance is lost to be led to exhaustion.

Therefore, after selecting a channel, if the user selects the all-channel reproduction mode, this is, if the selected channel is  
30 the right channel R, and if the all-channel reproduction mode is

selected, then the control section 120 controls the DAC 142 in such a manner that the signals of the right channel R are outputted through both the first and second speaker 150 and 152. In this manner, when using a headphone or speakers, the hearing balance can be maintained. This method is illustrated in FIG. 8.

FIG. 9 illustrates the learning method of the present invention in which multiple channels are used. That is, the stereo method is expanded, so that the respective channels can be subjected to the selections, and that the selected channel signals can be outputted through all the speakers.

After learning the talks of the talker A, the user can learn the talks of the talker B by turning off the talk intervals of the talker B, and by turning on the talk intervals of the talker A. Under this circumstance also, the caption data can be set in a selective manner.

In the above description, there were only two talkers. However, by providing multiple channels, the talks of the talkers A, B, C, D ... can be efficiently learned based on the above described principle.

In the case of karaoke, the songs and the melody accompaniments can be recorded in respective channels by adopting a two-channel method. In this case, the songs and the melody accompaniments are respectively of the mono type. If the songs and the melody accompaniments are made to be of stereo type, then at least four channels are required.

In the case where two channels are simultaneously reproduced, the songs and the melody accompaniments are separately outputted, and therefore, the user can learn the songs in an easy manner. Further, after making the songs somewhat familiar to the user, if the song channel is turned off,

then only the melody accompaniments are reproduced. Accordingly, the user can sing the songs while listening to the melody accompaniments. Further, after perfectly learning the songs, the user can turn on both of the channels to reproduce the songs and the melody accompaniments simultaneously, so that the user can sing the songs like a singer. Under this circumstance also, the caption data can be displayed. Further, by using multiple channels (more than two channels), a chorus or duet can be performed by selectively reproducing the multiple channels.

In the method of the present invention, not only the audio signals but also the caption data can be utilized. In other words, the caption data can be selectively displayed in relation to the audio signals, and in this manner, the difficulty level of the learning can be adjusted. That is, the caption data can be turned on or off in accordance with the learning progress. When the user memorizes all the talks, all the caption data are kept from being displayed, and only the sequence such as A and B is displayed, so that the rest of the text is trusted to the memory of the user in carrying out the conversation.

Further, in the learning apparatus of the present invention, the following functions are provided. That is, the learning apparatus for learning by using an output channel selection for a caption data according to the present invention is characterized in that: If an operation-on signal for reproducing the audio signals from a keypad 180 is an input, an operation mode during a reproduction which is currently set by a control section 120 is checked; if the operation mode is normal, the control section 120 controls a decoder 130 to output the audio signals which have been set to respective channels (R and L); if the operation mode is set to the right channel (R), the control section controls the decoder



130 to reproduce and output the audio signals to the right channel (R); and if the operation mode is set to the left channel (L), the control section controls the decoder to reproduce and output the audio signals to the left channel. That is, the user can exercise  
5 the selections as defined above, and the apparatus for making the above operation possible should come within the scope of the present invention.

In the apparatus having the above described functions, not only the speakers but also the picture display devices can be  
10 added as many as required, so that the audio signals can be linked to the caption data.

The present invention is applicable not only to the language learning apparatus and the karaoke but also to the conventional personal computers. FIG. 10 illustrates the structure of the  
15 conventional personal computer. If this computer is compared with FIGs. 5 and 6, the role of the decoder 130 can be realized by program in the CPU+MB. Further, the reproduction of audio signals can be realized by a sound card and speakers. The picture display device can be embodied by the graphic card and a monitor.  
20 The digital files can be stored in HDD or in CD, and therefore, the computer has the functions equivalent to those of the language learning apparatus or the karaoke. Thus the objects of the present invention can be accomplished through the conventional computers.

According to the present invention as described above,  
25 when using the digital files for learning a foreign language or songs, the learning can be carried out in an arbitrary manner, thereby improving the efficiency of the learning.

That is, the user can arbitrarily adjust the progress of the learning in accordance with the level of the achieved learning.

**WHAT IS CLAIMED IS:**

1. A method for learning by using a digital audio and its caption data, comprising the steps of:

5 forming a first learning pattern storing mode for storing a song caption, a voice of an original singer, and a melody accompaniment by converting their signals into a digital file; and forming a second learning pattern storing mode for storing a song caption and a melody accompaniment by converting their signals into a digital file,

10 whereby a digital file is used for an arbitrary song, and the digital file is reproduced by using the first or second learning pattern storing mode so as to facilitate learning an arbitrary song.

2. A storing method for storing components of songs by utilizing a digital audio, comprising the steps of:

15 forming a first learning pattern storing mode for storing a song caption, a voice of an original singer, and a melody accompaniment by converting their signals into a digital file;

forming a second learning pattern storing mode for storing a song caption and a melody accompaniment by converting their signals into a digital file in respectively storable forms; and

20

forming a third learning pattern storing mode for storing a melody accompaniment by converting their signals into a digital file,

25 whereby one or two or more of the above storing modes are combined to store the components of the song.

3. A method for learning by using a digital audio and its caption data, comprising the steps of:

forming a first learning pattern storing mode for storing a voice and a caption of a foreign language speech or a news by distinguishing the voice of a speaker and the caption of speech details in letters or news details in letters, and by converting  
5 signals of the audio and caption to a digital file; and

forming a second learning pattern storing mode for storing only a voice of a foreign language speech or a news by distinguishing the voice of a speaker and the caption of speech details in letters or news details in letters, and by converting  
10 signals of only the voice of the speaker to a digital file,

whereby a digital file is formed for an arbitrary speech or news, and the digital file is reproduced in accordance with a selection of a reproduction by a user so as to make it possible to arbitrarily learn a language through the speech or news.

15 4. The method as claimed in claim 3, further comprising the step of: forming a third learning pattern storing mode for storing talkers' voices, the caption of the speech or news, and a translation of the speech or news in a form of a digital file.

20 5. A storing method for storing a speech or news, comprising the steps of:

forming a first learning pattern storing mode for storing a voice and a caption of a foreign language speech or a news by distinguishing the voice of a speaker and a caption of speech details in letters or news details in letters, and by converting  
25 signals of the audio and caption to a digital file; and

forming a second learning pattern storing mode for storing only the voice of a foreign language speech or a news by distinguishing the voice of a speaker and a caption of speech

details in letters or news details in letters, and by converting signals of only the voice of the speaker to a digital file,

whereby an arbitrary speech or news is stored in a form of a digital file.

5           6. A method for learning by using a digital audio and its caption data, comprising the steps of:

          forming a first learning pattern storing mode for recording a full sound - full caption by preparing a digital data file of all voices and all talk captions of all talkers of a foreign movie or  
10       drama; and

          forming a second learning pattern storing mode for storing a voice of a data file by recording a scenario of the movie or drama after deleting voices of certain talkers so as to make a user talk in place of the deleted voices,

15       whereby a digital data is formed, and if the user selects a learning reproduction mode and selects the talkers, the digital data file is selectively reproduced so as to make the user talk in place of the particular talkers.

20       7. The method as claimed in claim 6, wherein when inputting names of talkers and the caption data, serial codes are assigned for the talkers instead of the names of the talkers, and the names of the talkers are respectively matched to the serial codes, whereby the captions and audio outputs of particular talkers can be selectively deleted when carrying out the learning.

25       8. The method as claimed in claim 6, wherein the digital file prepared by the first and second learning pattern storing modes can be transmitted through a wire such as a printer port

(parallel port), a serial port, USB (universal serial bus), firewire (IEEE 1394), or through a wireless route such as an infrared ray data or a blue tooth.

5 9. The method as claimed in claim 7, wherein the digital file storing means of a reproduction apparatus is a non-volatile memory such as flash memory, or a DVD (digital versatile disk).

10 10. A method for learning by using an output channel selection for a caption data and an audio, comprising the steps of:  
checking an operation mode of a current reproduction operation upon inputting an operation-on signal by a user for reproducing audio signals (first step);

outputting audio signals which have been set to respective channels (R and L), if the operation mode is found to be a normal channel outputting (second step);

15 reproducing and outputting the audio signals to the right channel if the operation mode is set to the right channel (R) (third step); and

20 reproducing and outputting the audio signals to the left channel (L) if the operation mode is set to the left channel (fourth step).

25 11. The method as claimed in claim 10, wherein at the third and/or fourth step, when reproducing the selected channel output signals, the signals of the selected channel are outputted also through non-selected channels so as make the selected channel signals outputted through two channels (R and L).

12. The method as claimed in anyone of claims 10 and 11,

wherein the caption data is outputted in synchronization with the output of the audio signals of the selected channel.

13. The method as claimed in claim 12, wherein the caption data synchronized with the audio signals can be turned on or off in accordance with a progress degree of the learning, a difficulty level, or an individual's taste.

14. The method as claimed in claim 10, wherein the digital file can be transmitted through a wire such as a printer port (parallel port), a serial port, USB (universal serial bus), firewire (IEEE 1394), or through a wireless route such as an infrared ray data or a blue tooth.

15. The method as claimed in claim 10, wherein the digital file storing means of a reproduction apparatus is a non-volatile memory such as flash memory, or a DVD (digital versatile disk).

16. A method for learning by using an output channel selection for a caption data and an audio by using three or more channels, comprising the steps of:

checking an operation mode of a current reproduction operation upon inputting an operation-on signal by a user for reproducing audio signals (first step);

outputting audio signals which have been set to respective channels (R and L), if the operation mode is found to be a normal channel outputting (second step); and

reproducing and outputting the signals to a particular channel if the operation mode is set to the particular channel (R) (third step).

17. The method as claimed in claim 16, wherein at the third step, when reproducing the selected channel output signals, the signals of the selected channel are outputted also through non-selected channels so as make the selected channel signals  
5 outputted also through the rest of the channels.

18. The method as claimed in claim 17, wherein the caption data is outputted through a display screen of a reproduction apparatus in synchronization with the output of the audio signals of the selected channel.

10 19. The method as claimed in claim 18, wherein the caption data synchronized with the audio signals can be turned on or off in accordance with a progress of the learning, a difficulty level, or an individual's taste.

15 20. A learning apparatus for learning by exercising an output channel selection, characterized in that:

if an operation-on signal for reproducing audio signals from a keypad is an input, an operation mode during a reproduction which is currently set by a control section is checked;

20 if the operation mode is normal, the control section controls a decoder to output the audio signals which have been set to respective channels (R and L);

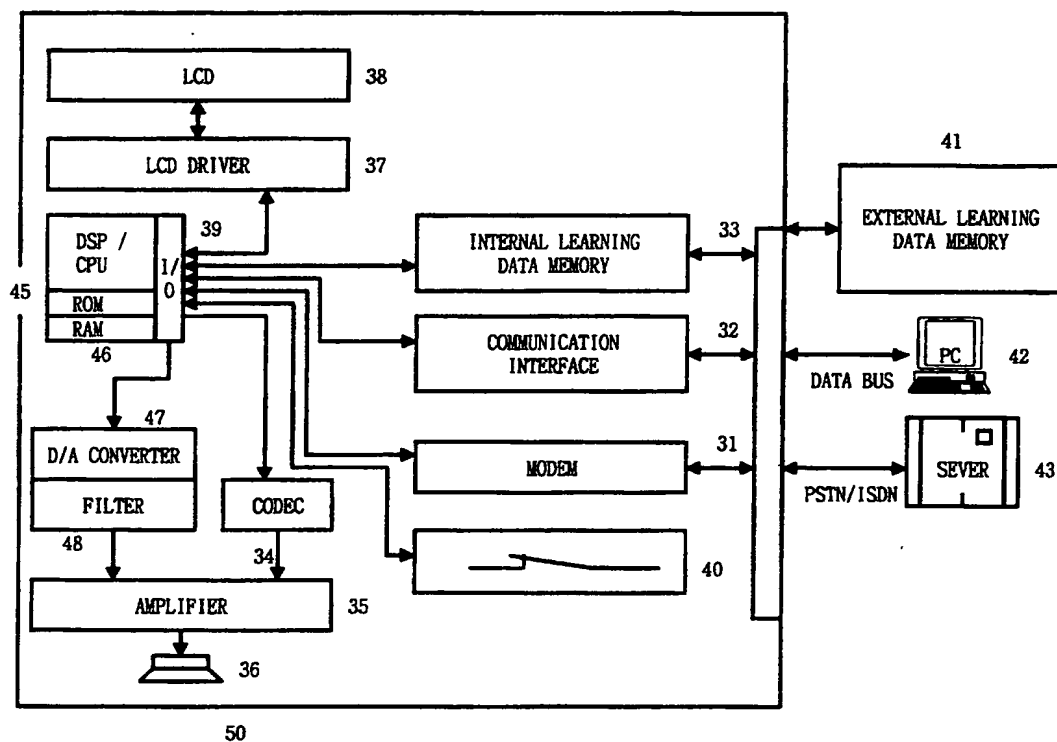
if the operation mode is set to the right channel (R), the control section controls the decoder to reproduce and output the audio signals to the right channel; and

25 if the operation mode is set to the left channel (L), the control section controls the decoder to reproduce and output the audio signals to the left channel.

21. The learning apparatus as claimed in claim 20, wherein the caption data is outputted through a display screen of a reproduction apparatus in synchronization with the output of the audio signals of the selected channel.

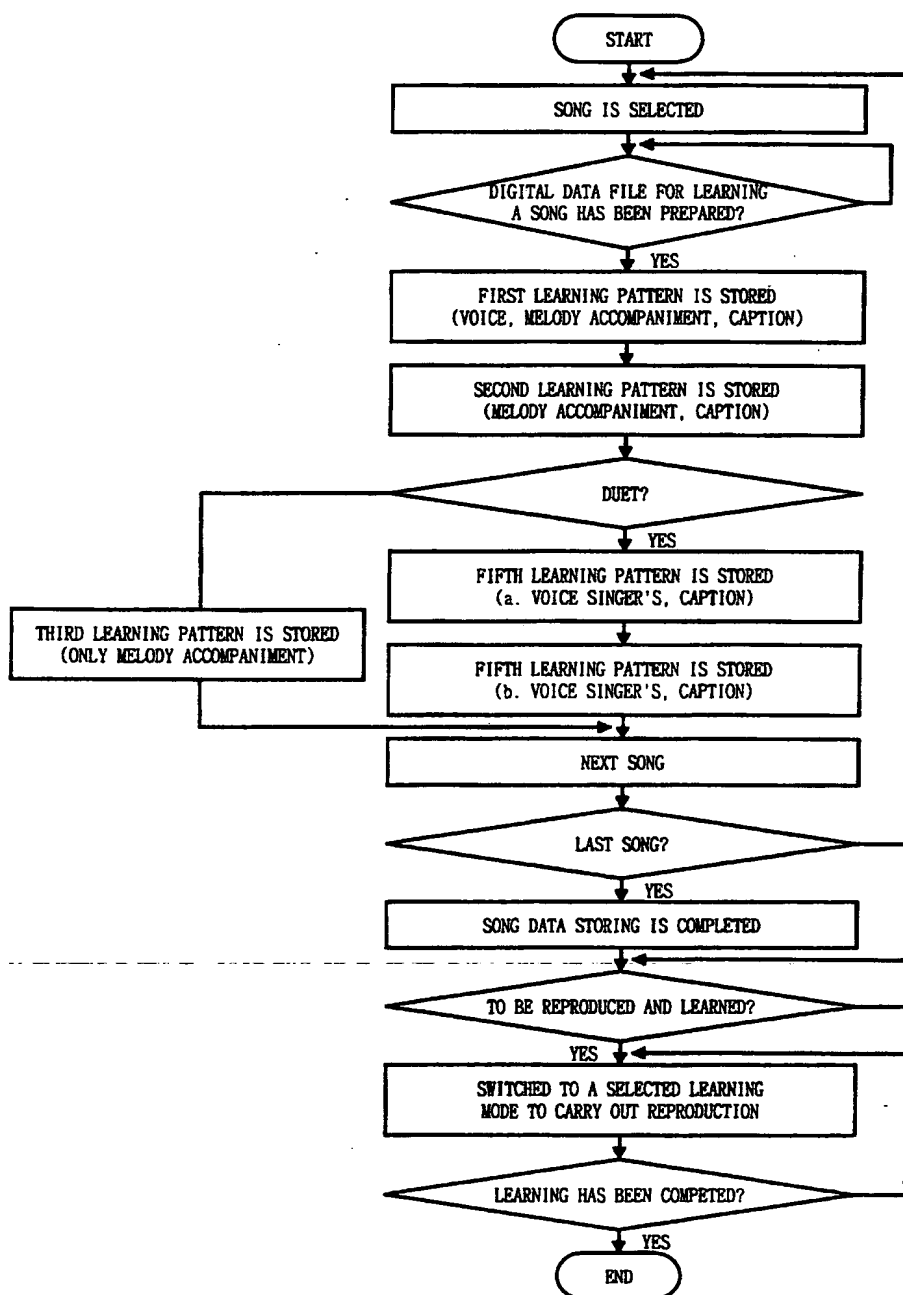


1/12  
FIG. 1



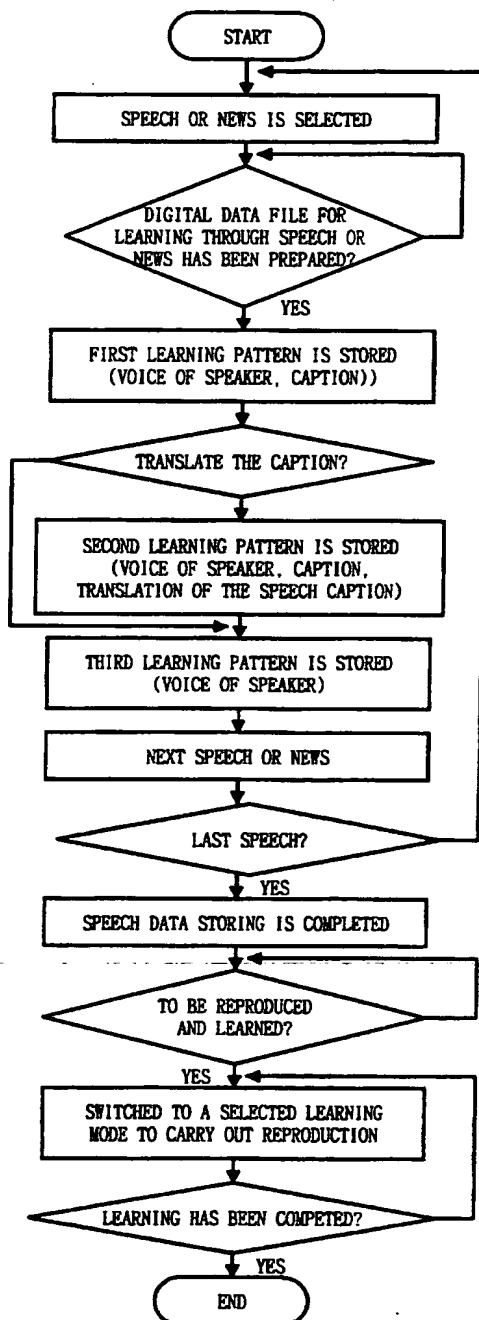
**THIS PAGE BLANK (USPTO)**

2/12  
FIG. 2



**THIS PAGE BLANK (USPTO)**

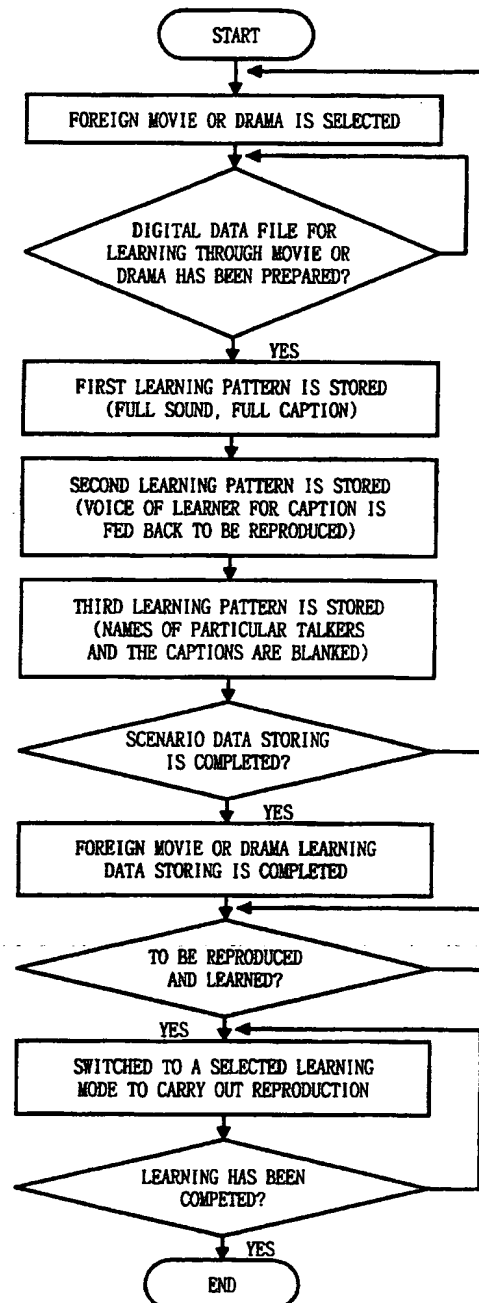
3/12  
FIG. 3



**THIS PAGE BLANK (USPTO)**

4/12

FIG. 4



**THIS PAGE BLANK (USPTO)**



5/12  
FIG. 5a

Andy : How are you?

\_\_\_\_\_ :

brother : I am fine.

**THIS PAGE BLANK (USPTO)**

6/12  
FIG. 5b

Andy : How are you?

----- :

brother : I am fine.

**THIS PAGE BLANK (USPTO)**

7/12  
FIG. 5c

Andy : \_\_\_\_\_

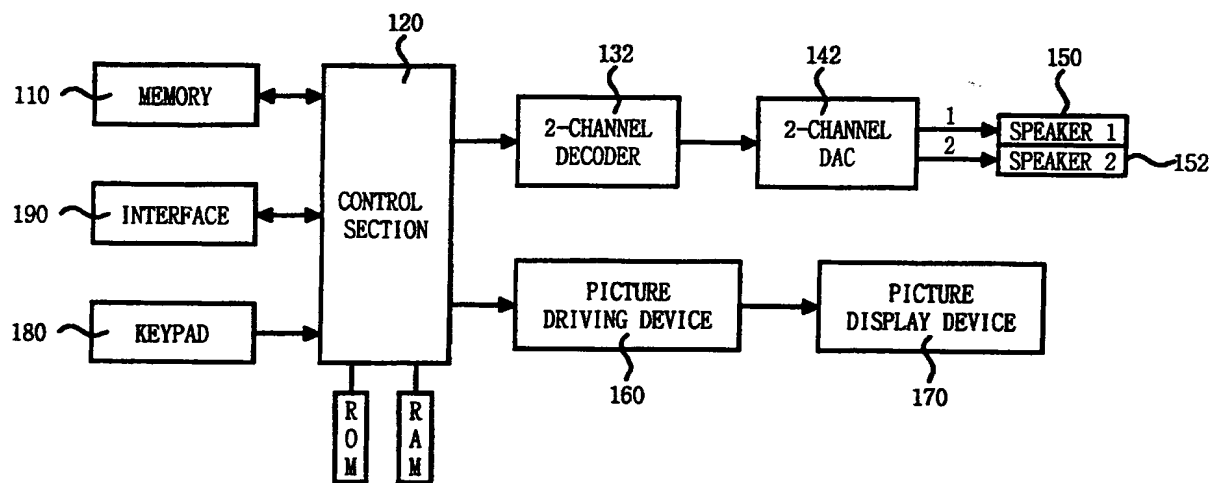
\_\_\_\_\_ : \_\_\_\_\_

brother : I am fine.

THIS PAGE BLANK OF

8/12

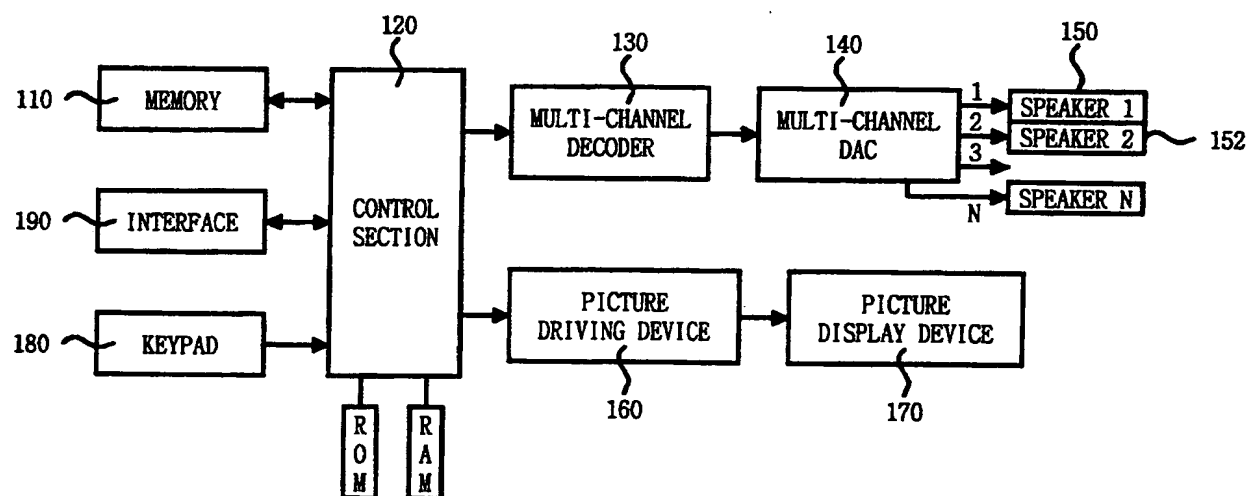
FIG. 6



**THIS PAGE BLANK (uspto)**



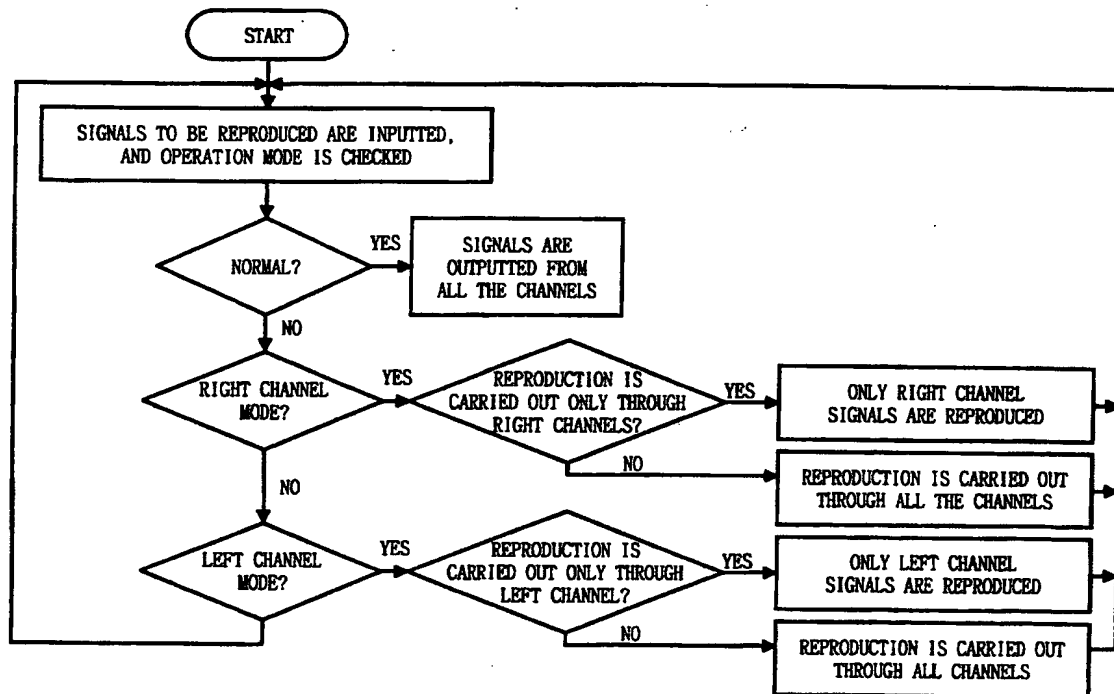
9/12  
FIG. 7



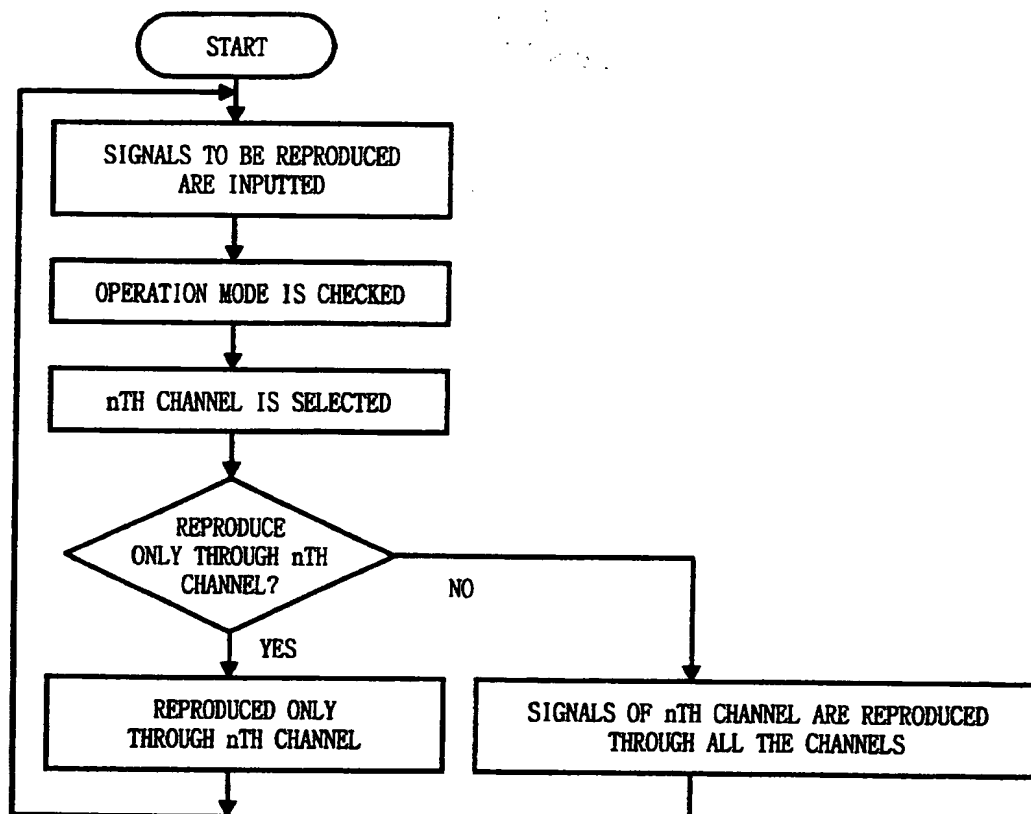
**THIS PAGE BLANK (USPTO)**

10/12

FIG. 8



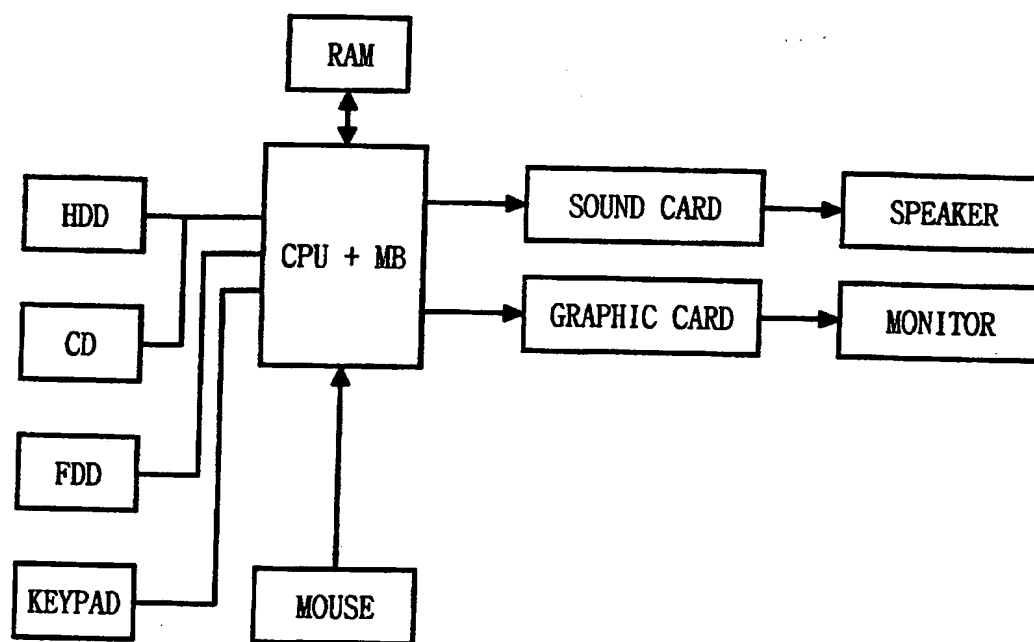
**THIS PAGE BLANK (USPTO)**

11/12  
FIG. 9

**THIS PAGE BLANK (uspto)**

12/12

FIG. 10



THIS PAGE BLANK (USPTO)



## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/KR00/00836

**A. CLASSIFICATION OF SUBJECT MATTER****IPC7 G06F 17/60**

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC7 G06F17/60, G11B, H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Korean Patent and applications for inventions since 1975

Japanese Patents and applications for inventions since 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PAJ, USPATFULL, WPI

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 6097442 A(TOMPSON CONSUMER ELECTRONICS INC.) 1 AUG 2000	1-21
A	US 5999497 A (SONY CORPORATION) 7 DEC 1999	1-21
A	US 5572260 A (MITSUBISHI ELECTRIC SEMICONDUCTOR SOFTWARE CO.) 5 NOV 1996	1-21
A	US 5543850 A (CIRRUS LOGIC INC.) 6 AUG 1996	1-21
A	US 4115930 A (BECK CHARLES R.) 26 SEPT 1978	1-9



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

07 NOVEMBER 2000 (07.11.2000)

Date of mailing of the international search report

08 NOVEMBER 2000 (08.11.2000)

Name and mailing address of the ISA/KR

Korean Industrial Property Office  
Government Complex-Taejon, Dunsan-dong, So-ku, Taejon  
Metropolitan City 302-701, Republic of Korea

Facsimile No. 82-42-472-7140

Authorized officer

JEON, Hyun Jin

Telephone No. 82-42-481-5788



**THIS PAGE BLANK (uspto)**